Holistic Computation Operational Blueprint (4Capital × Causal × ML × 4Es)

Version: v1.0
Date: Sept 3, 2025
Owner/Author: Dr. Alex Liu / RMDS Lab
Applies to: Research, Data Science, Decision Science, and MLOps teams

---

A step-by-step, production-ready blueprint to turn predictive modeling into **decision systems** that deliver value across **4Capital** (Material, Intellectual, Social, Spiritual), grounded in **causal science**, **machine learning**, and the **4Es** (Equation → Estimation → Evaluation → Execution).

Quick Start (What you will produce)

- **4Capital OKRs** + **Utility & Constraint spec**
- **Causal DAG** + **estimand definitions**
- **Cohort & data contracts** + **feature store spec**
- **Models** (forecasting + uplift/causal ML) with **calibration & uncertainty**
- **Policy** (mapping predictions → actions) + **simulation results**
- **Evaluation report** (predictive + policy value + fairness + robustness)
- **Monitoring plan** + **Model & Policy cards**
- **Reusable pipelines & playbooks**

Roles & RACI (suggested)

Activity	Sponsor	PI/ Lead	Causal Lead	Data Eng	ML Eng	Product/ Ops	Compliance
4Capital goals	A	R	C	C	C	C	C
DAG & estimands	C	A	R	C	C	C	C
Data contracts	C	A	C	R	C	C	C
Modeling (4Es)	C	A	R	C	R	C	C
Simulation	C	A	C	C	R	R	C
Evaluation & policy	C	A	R	C	R	R	C

Activity	Sponsor	PI/ Lead	Causal Lead	Data Eng	ML Eng	Product/ Ops	Compliance
Governance/monitoring	C	A	C	C	R	R	R
Legend: R = Responsible, A = Accountable, C = Consulted							

Phase 1 — Set Goals with the 4Capital Lens (What to optimize)

Objective: Define value, constraints, decision points, and population.

Tasks 1. Select **1 north-star KPI (Material)** and **2 companion KPIs** spanning **Intellectual/Social/Spiritual**. 2. Specify **decision points** (who/what/when), **population**, **time horizon**, and **capacity/operational constraints**. 3. Draft **utility function** and **risk/constraint register** (fairness, safety, compliance, reliability/ emissions, budget).

Artifacts - 4Capital OKRs
- Utility skeleton U(a, x)
- Constraint register & decision inventory

Acceptance Criteria - KPIs are **measurable**, **auditable**, and have **baselines & targets**
- Constraints quantified (thresholds)
- Decision inventory covers actor, timing, and allowable actions

Template — 4Capital OKRs | Capital | Objective | KPI | Baseline | Target | Timeframe | |---|---|---|---|---|---| | Material | | | | | | Intellectual | | | | | | Social | | | | | | Spiritual | | | | | |

Phase 2 — Causal Map & Model/Estimand Specification

Objective: Make assumptions explicit and choose what to estimate.

Tasks 1. Draft **DAG**: treatments/exposures, outcomes, confounders, mediators, instruments; note **SUTVA**, **positivity**, **no interference** assumptions. 2. Determine **adjustment set** (back-door/front-door), run **identifiability checks**. 3. Define **estimands** (ATE/CATE/ITE, uplift, structural equations) and the **policy variable** (what action will change). 4. Create **Assumptions Log** and **test plan** (placebo tests, sensitivity analyses).

Artifacts - DAG (image + machine-readable form)
- Minimal adjustment set + variable roles table
- Estimand specs & policy mapping diagram
- Assumptions Log

Acceptance Criteria - Adjustment set defensible; identifiability stated or limitations documented
- Estimands match decisions; policy variable operationalized

Phase 3 — Data Preparation aligned to the DAG

Objective: Build a leak-free, auditable dataset mapped to causal roles.

Tasks 1. Define **cohort** (inclusion/exclusion), **index date**, **lookback/forecast windows**, and **causal time order**. 2. Implement **leakage guards** (time-based splits, post-outcome data exclusion, frozen features at decision time). 3. Audit **missingness** (MCAR/MAR/MNAR), apply imputation/indicator strategies; document **proxies** and **lineage**. 4. Create **feature store** spec aligned to roles (treatment, confounder, mediator, instrument, outcome, eligibility). 5. Establish **data contracts** (schemas, freshness, quality checks) and **privacy/fairness attributes**.

Artifacts - Cohort spec, data dictionary, lineage diagram
- Missingness/bias coverage report
- Train/validation/test **temporal** splits + **policy-eval folds**
- Feature store definitions & data contracts

Acceptance Criteria - Leakage tests pass; no feature uses post-decision info
- Coverage $\geq$ threshold per key subgroups; missingness treatment documented
- Reproducible dataset build with versioned snapshots

Phase 4 — The 4Es Workflow

4E-1. Equation (Formalize the decision problem)

- Specify **utility**: $\max_{\pi \in \Pi} \mathbb{E}[U(A, X)]$ subject to **constraints** $g_j(A, X) \leq c_j$.
- Define **policy class** Π (threshold rules, prioritization, knapsack/queueing, bandits), **cost matrix**, and **fairness constraint or penalty**.
- Map **predictions → actions** (triage thresholds, treatment assignment, ranking, routing).
- Write pseudocode for the policy.

Output: Utility & constraint spec; policy pseudocode.

4E-2. Estimation (Fit models & quantify uncertainty)

- **Baselines**: transparent GLM/GBM; **Candidates**: RF/GBM/linear nets; **Causal ML**: T-/S-/X-/R-/DR-learners, causal forests, uplift trees, doubly robust learners.
- **Calibration**: isotonic/Platt; **Uncertainty**: conformal prediction, quantile regression, bootstrapping.
- **Cross-fitting**, **regularization**, **early stopping**; stable feature transforms.
- Explainability: global (permutation/SHAP) and local at decision boundary.

Output: Fitted models, calibration curves, uncertainty intervals, explainers.

4E-3. Evaluation (Predictive → Policy value → Robustness)

- **Predictive**: AUC/PR, Brier, RMSE/MAE, calibration error.
- **Policy value**: expected utility, cost–benefit, **uplift** metrics (Qini/uplift-AUC), **capacity utilization**, **constraint hit rate**.
- **Off-policy**: IPS/IPW, DR, Switch-DR, SNIPS; coverage diagnostics.
- **Fairness**: demographic parity diff, equalized odds, calibration-in-the-small, predictive parity; **recourse feasibility**.
- **Robustness**: temporal holdouts, covariate shift stress tests, sensitivity to unobserved confounding (e.g., Rosenbaum bounds), partial identification.

Output: Evaluation report with a **go/no-go** policy recommendation.

4E-4. Execution (Operationalize with guardrails)

- Choose thresholds/assignments via **utility under constraints** (grid search or Bayesian optimization).
- Implement **guardrails**: eligibility checks, business rules, human-in-the-loop, rate limiters, capacity-aware scheduling/queueing.
- Integration plan: APIs, feature store, batch/stream inference, **observability hooks**.
- Rollout design: shadow → limited A/B → staged expansion with **rollback**.

Output: Policy card, runbooks, integration checklist, SLA.

Phase 5 — Policy Simulation & Scenario Analysis

Objective: Validate decisions before they affect people or operations.

Tasks 1. Build a **simulation notebook** that ingests predictions, policy, constraints, and capacity.
2. Generate **scenarios**: demand spikes, mix shifts, delays, outages, price/emissions changes; include worst-cases.
3. Use **Monte Carlo** or **agent-based** modeling; model service-time & arrival distributions; include queueing.
4. Run **off-policy evaluation** using logged data; verify overlap/positivity; report confidence intervals.
5. Optimize thresholds/dispatch to maximize utility subject to constraints.

Artifacts - Scenario pack & parameter sheet
- Threshold selection table
- KPI dashboard & constraint heatmaps
- Decision logs for audit

Acceptance Criteria - Chosen policy meets KPI targets with **< defined** constraint violation rates across scenarios
- Sensitivity analysis documented; key risks mitigated

Phase 6 — Governance, Monitoring & Learning Loop

Objective: Keep the system safe, fair, stable, and improvable.

Monitoring Plan - **Performance**: KPI deltas vs. baseline, calibration drift, prediction/decision latency.
- **Fairness**: subgroup metrics, equalized odds gaps, false positive/negative parity, recourse success rates.
- **Drift & Stability**: PSI/JS-divergence, concept drift tests, feature/label freshness.
- **Uncertainty**: prediction interval coverage; abstention/deferral rates.
- **4Capital**: Material ROI, Intellectual asset accrual (playbooks, reusable code), Social metrics (access, parity), Spiritual metrics (e.g., verified emissions, ethics compliance).

Ops & Compliance - **Auditability**: end-to-end lineage (data → model → policy → decision), reproducible runs (seeds, environments), signed artifacts.
- **Controls**: approvals, change logs, access policies, privacy & sector rules mapping (HIPAA, market ops, privilege rules, etc.).
- **Recourse**: user-facing explanations, appeal pathways, override capture.

Triggers - Retraining/rollback criteria; incident response runbook; periodic ethics review.

Artifacts - Monitoring dashboards, audit trail, change control records, post-incident reviews.

Phase 7 — Scale & Reuse (Responsible scaling)

Objective: Make the practice durable and portable.

Tasks 1. Package **pipelines** (data → model → policy → simulation → monitoring) with templates and IaC.
2. Establish **feature store** & **model/policy registry**; enforce **data contracts**.
3. Create a **knowledge base**: DAG worksheets, assumptions logs, model/policy cards, evaluation playbooks.
4. Run **quarterly portfolio reviews** across the 4Capital metrics; sunset/refresh underperforming policies.
5. Maintain **ethical risk register** and **stakeholder feedback loops**.

Artifacts - Reusable libraries, scaffolds, CI/CD & MLOps configs
- Program-level 4Capital scorecards
- Training materials & onboarding kit

Key Templates (copy/paste)

A. Assumptions Log

ID	Assumption	Rationale	Test/Probe	Status	Owner
A1					

B. Variable Roles Table

Variable	Role (T/C/M/I/O/E)	Source	Time	Quality	Notes

C. Utility & Constraints (spec)

- **Utility:**
  $U(a, x) = w_M \cdot M(a, x) + w_I \cdot I(a, x) + w_S \cdot S(a, x) + w_P \cdot P(a, x) - \mathrm{Cost}(a, x)$
- **Constraints:**
  $g_1(a, x) \le c_1$ (capacity), $g_2(a, x) \le c_2$ (emissions), fairness constraint $|\mathrm{gap}_{group}| \le \delta$.

D. Policy Card (minimum contents)

- Purpose & scope; decision boundary; inputs & freshness
- Mapping from scores to actions; thresholds & overrides
- Guardrails & eligibility; fairness and recourse commitments
- Calibration & uncertainty treatment; monitoring & triggers
- Owners; change history; validation summary

E. Evaluation Report (minimum contents)

- Dataset & cohort; leakage checks; splits
- Model family & hyperparams; calibration & uncertainty
- Predictive metrics; **policy value** metrics; capacity & constraint hits
- Fairness & stability results; sensitivity analyses
- Limitations; go/no-go recommendation

F. Monitoring Checklist

- Metrics tracked (performance, fairness, drift, uncertainty, 4Capital)
- Alert thresholds & actions; retraining cadence
- Data & model versioning; dashboard links
- Incident response steps; communication template

Risk Register (examples)

- **Data leakage** → Temporal splits; decision-time feature freeze; audits.
- **Unobserved confounding** → Sensitivity analyses; encourage experimental data; IVs when valid.
- **Fairness regressions** → Pre/post-deployment subgroup monitoring; mitigation with constraints or post-processing.
- **Capacity misestimation** → Simulation with queueing; safety buffers; staged rollout.
- **Concept drift** → PSI & calibration drift alarms; fallback policy; canary tests.
- **Explainer misuse** → Decision-boundary explanations with uncertainty & limitations.
- **Compliance gaps** → Control mapping; periodic audits; documented recourse.

Glossary (select)

- **4Capital:** Material, Intellectual, Social, Spiritual value dimensions.
- **DAG:** Directed acyclic graph encoding causal assumptions.
- **Estimand:** Quantity to estimate (e.g., ATE, CATE, uplift).
- **Policy value:** Expected utility of mapping from predictions to actions.
- **Off-policy evaluation:** Estimating policy value from logged data without randomized trials.
- **Recourse:** User's ability to change inputs to affect decisions.

How to Use

- Treat each Phase as a **gate** with artifacts and acceptance criteria.
- Store artifacts in a versioned repo; link to dashboards and code.
- Revisit 4Capital OKRs quarterly; refresh the DAG when domain knowledge changes.